

# Estimation of linear observer templates in the presence of multi-peaked Gaussian noise through 2AFC experiments

Darrin C. Edwards\*, Matthew A. Kupinski, Robert M. Nishikawa, and Charles E. Metz

Kurt Rossmann Laboratories for Radiologic Image Research,  
Department of Radiology, The University of Chicago, Chicago IL 60637

## ABSTRACT

We extend a method for linear template estimation developed by Abbey *et al.* which demonstrated that a linear observer template can be estimated effectively through a two-alternative forced choice (2AFC) experiment, assuming the noise in the images is Gaussian, or multivariate normal (MVN). We relax this assumption, allowing the noise in the images to be drawn from a weighted sum of MVN distributions, which we call a multi-peaked MVN (MPMVN) distribution. Our motivation is that more complicated probability density functions might be approximated in general by such MPMVN distributions. Our extension of Abbey *et al.*'s method requires us to impose the additional constraint that the covariance matrices of the component peaks of the signal-present noise distribution all be equal, and that the covariance matrices of the component peaks of the signal-absent noise distribution all be equal (but different in general from the signal-present covariance matrices). Preliminary research shows that our generalized method is capable of producing unbiased estimates of linear observer templates in the presence of MPMVN noise under the stated assumptions. We believe this extension represents a next step toward the general treatment of arbitrary image noise distributions.

**Keywords:** Signal detection, observer template, model observer, computer-aided diagnosis, two-alternative forced choice detection, non-Gaussian noise model

## 1. INTRODUCTION

Computer-aided diagnosis (CAD) is a diagnosis made by a radiologist who makes use of computer analysis of patient information, such as a patient image. The goal of CAD is to improve the diagnostic accuracy and consistency of the radiologist<sup>1</sup> while avoiding some of the complexities entailed in obtaining and combining multiple human decisions.<sup>2-4</sup> One task of interest in the design of CAD schemes is the detection of signals, *i. e.*, deciding whether an image contains an abnormality or not.

Signal detection algorithms used in CAD schemes are typically complicated and highly nonlinear. This can make analysis of such algorithms — *e. g.*, optimizing their parameters nonstochastically — extremely difficult or impossible. Ideally, but quite impractically, we would require a closed-form expression for the CAD algorithm's mapping function from image space to decision variable space. Lacking this, a linear approximation to the mapping function might still be useful.

Recently Abbey *et al.* developed an elegant method of estimating a linear observer's template using only the decisions made by the observer in a two-alternative forced choice (2AFC) experiment.<sup>5</sup> The stated motivation for their work was to estimate human observer templates (in detection tasks in which the human observer behavior can be regarded as linear) for comparison with observer models; their results are quite general, however, and valid for any observer which uses a linear function of the image data to produce a decision variable. Indeed, the beauty of this method is the limited assumptions made about the observer — the observer is not even required to produce a “measurable” decision variable, only to choose between two images in each trial of a 2AFC experiment. The only assumption made regarding the images input to the observer in the 2AFC experiment is that the noise in both the signal-present and signal-absent images is Gaussian, or multivariate normal (MVN), with known means and covariance matrices.

Frequently, however, one must allow for multiple sources of noise in images, for example photon noise and background-structure noise in a medical image. Such a situation cannot be adequately described in general by a “single-peak” MVN distribution. It is then natural to ask whether the template estimation method of Abbey *et al.*

---

\*Correspondence: E-mail: d-edwards@uchicago.edu; Telephone: 773 834 5094; Fax: 773 702 0371

can be extended to image noise probability density functions (PDFs) which are more complicated than MVN. Such an extension would be extremely difficult in general, however, because the success of their method depends crucially on the attractive properties of the MVN distribution; *e. g.*, a linear combination of MVN random variables is again MVN. Applying this method to images with Poisson distributed noise, for example, would require expressions for the PDFs of arbitrary linear combinations of Poisson random variables.

One situation to which the method of Abbey *et al.* does appear to be extendable is that in which the image noise PDF is a weighted sum of MVN functions. We will refer to this PDF as a multi-peaked MVN (MPMVN) distribution. A random variable with such a distribution may be thought of intuitively as conditionally MVN, conditional on the outcome of a discrete random variable. A more physical example would be an imaging system with multiple sources of noise, such as the film granularity and x-ray quantum components of noise in radiographs.

This extension to the MPMVN case is not completely general, because it requires that all the component MVN covariance matrices in the signal-present PDF be equal, and that all the component MVN covariance matrices in the signal-absent PDF be equal. One of our motivations for exploring the MPMVN extension to the method of Abbey *et al.*, in addition to its tractability, is the hope that other, practically relevant PDFs might be approximated by sums of very narrow MVN peaks. For such an application, the limitation stated might be acceptable.

In this paper, we show how the template estimation method of Abbey *et al.* can be extended to images for which the signal-present and signal-absent image PDFs are both MPMVN. In section 2.1, we review the theoretical results of Abbey *et al.* and extend them to the MPMVN case. Section 2.2 describes the simulation studies we performed to show that our extension produces unbiased estimates of linear observer templates. The results of these studies are presented in section 3. In sections 4 and 5 we discuss these results and our conclusions from this preliminary investigation.

## 2. MATERIALS AND METHOD

### 2.1. Theory

#### 2.1.1. Method of Abbey *et al.*

In the framework outlined in Ref. 5, images (represented by vectors of graylevel values) are drawn from two distributions,

$$\vec{g}^+ \sim \text{MVN}(\vec{s} + \vec{b}, \Sigma^+) \quad \text{and} \quad (1)$$

$$\vec{g}^- \sim \text{MVN}(\vec{b}, \Sigma^-), \quad (2)$$

where the images  $\vec{g}^+$  are referred to as “signal-present”, and the images  $\vec{g}^-$  are referred to as “signal-absent”. (Throughout this paper random variables are denoted in boldface type, vector quantities by raised arrows, and estimators by circumflexes.) Here  $\vec{s}$  is the signal of interest,  $\vec{b}$  is the nonrandom component of the background, and  $\Sigma^+$  and  $\Sigma^-$  are the covariance matrices of the signal-present and signal-absent distributions. Also, throughout this paper “graylevel values” should be understood as abstract quantities ranging from  $-\infty$  to  $+\infty$ , and not the finite range of discrete values to which the graylevel values must be scaled in order to produce a displayable image.

A linear observer forms a decision variable by applying a template  $\vec{t}$  to an image. The observer’s decision variables will then come from two distributions, defined *via*

$$\mathbf{r}^+ \equiv \vec{t}^\dagger \vec{g}^+ + \epsilon^+ \quad \text{and} \quad (3)$$

$$\mathbf{r}^- \equiv \vec{t}^\dagger \vec{g}^- + \epsilon^-. \quad (4)$$

Here  $\epsilon^+$  and  $\epsilon^-$  represent the observer’s internal noise for the signal-present and signal-absent images, respectively. In Ref. 5, the task of interest is to estimate the template used by a human observer, and the internal noise is modeled as

$$\epsilon^+, \epsilon^- \sim \text{N}(0, \sigma_\epsilon^2). \quad (5)$$

We are concerned here primarily with deterministic algorithms for signal detection, and so for simplicity we will take the internal noise to be identically zero in our final results. We retain it as defined above in all derivations only for generality.

In a 2AFC trial, the observer is shown a signal-present image  $\vec{g}^+$  and a signal-absent image  $\vec{g}^-$ , and then chooses the image believed to contain the signal by choosing the greater of the decision variables  $r^+$  and  $r^-$ . If we ignore the case  $r^+ = r^-$ , irrelevant for continuously distributed random variables, one can write

$$\text{decision correct if } r^+ > r^-, \quad (6)$$

$$\text{decision incorrect if } r^+ < r^-. \quad (7)$$

This can be rewritten in terms of the image and observer variables as

$$\text{decision correct if } \vec{t}^\dagger \Delta \vec{g} + \Delta \epsilon > 0, \quad (8)$$

$$\text{decision incorrect if } \vec{t}^\dagger \Delta \vec{g} + \Delta \epsilon < 0, \quad (9)$$

where  $\Delta \vec{g} \equiv \vec{g}^+ - \vec{g}^-$  and  $\Delta \epsilon \equiv \epsilon^+ - \epsilon^-$ .

If the image chosen by the observer is denoted by  $\vec{g}^{\text{chosen}}$ , and the image rejected by  $\vec{g}^{\text{rejected}}$ , then the difference  $\vec{g}^{\text{chosen}} - \vec{g}^{\text{rejected}}$  is clearly  $\pm \Delta \vec{g}$ , the sign determined by whether a correct or incorrect decision is made. This observation leads immediately to Eqn. 4 of Ref. 5 which describes the outcome of a general 2AFC trial for a linear observer:

$$\vec{g}^{\text{chosen}} - \vec{g}^{\text{rejected}} = \Delta \vec{g} \text{sgn}(\vec{t}^\dagger \Delta \vec{g} + \Delta \epsilon). \quad (10)$$

Here  $\text{sgn}(\cdot)$  is the signum function, defined as

$$\text{sgn}(x) = \begin{cases} -1 & : x < 0 \\ 0 & : x = 0 \\ 1 & : x > 0. \end{cases} \quad (11)$$

A 2AFC experiment is then just a sequence of 2AFC trials and their outcomes, for a given observer and sequence of image pairs. The number of trials in the experiment is denoted by  $N_T$ .

Given the definitions in Eqns. 1, 2, and 5, the PDFs of the dependent random variables in Eqn. 10 are readily determined to be

$$\Delta \vec{g} \sim \text{MVN}(\vec{s}, \Sigma^+ + \Sigma^-) \quad \text{and} \quad (12)$$

$$\Delta \epsilon \sim \text{N}(0, 2\sigma_\epsilon^2). \quad (13)$$

Abbey *et al.* show<sup>5</sup> that, given these PDFs, the expectation value of Eqn. 10 with respect to both  $\Delta \vec{g}$  and  $\Delta \epsilon$  can be evaluated explicitly. The result, their Eqn. 20, is (allowing for minor changes in notation)

$$\langle \vec{g}^{\text{chosen}} - \vec{g}^{\text{rejected}} \rangle = (2P_C - 1)\vec{s} + \frac{2}{\sqrt{2\pi(2\sigma_\epsilon^2 + \vec{t}^\dagger(\Sigma^+ + \Sigma^-)\vec{t})}} \exp\left(\frac{-(\vec{t}^\dagger \vec{s})^2}{2(2\sigma_\epsilon^2 + \vec{t}^\dagger(\Sigma^+ + \Sigma^-)\vec{t})}\right) (\Sigma^+ + \Sigma^-)\vec{t}. \quad (14)$$

Here  $P_C$  is the expected value of the proportion of 2AFC trials in which the observer makes a correct decision. This is also evaluated explicitly by Abbey *et al.* to obtain

$$P_C = \Phi\left(\frac{\vec{t}^\dagger \vec{s}}{\sqrt{2\sigma_\epsilon^2 + \vec{t}^\dagger(\Sigma^+ + \Sigma^-)\vec{t}}}\right) \quad (15)$$

where  $\Phi(\cdot)$  is the standard normal cumulative distribution function.

Now by defining

$$\vec{z} \equiv \langle \vec{g}^{\text{chosen}} - \vec{g}^{\text{rejected}} \rangle \quad (16)$$

and

$$\alpha \equiv \frac{2}{\sqrt{2\pi(2\sigma_\epsilon^2 + \vec{t}^\dagger(\Sigma^+ + \Sigma^-)\vec{t})}} \exp\left(\frac{-(\vec{t}^\dagger \vec{s})^2}{2(2\sigma_\epsilon^2 + \vec{t}^\dagger(\Sigma^+ + \Sigma^-)\vec{t})}\right), \quad (17)$$

Eqn. 14 can be rewritten as

$$(\Sigma^+ + \Sigma^-)^{-1}(\vec{z} - (2P_C - 1)\vec{s}) = \alpha\vec{t}. \quad (18)$$

Note that  $\vec{t}$  appears explicitly on the right hand side, multiplied by the factor  $\alpha$ . Although in principle  $\alpha$  also depends on the observer template  $\vec{t}$ , the direction of  $\vec{t}$  in image space is of greater relevance, and we can consider determination of the quantity  $\alpha\vec{t}$  as equivalent to determining  $\vec{t}$ . (Abbey *et al.* note that changing scale from  $\vec{t}$  to  $c\vec{t}$  for some constant  $c$  has no effect if the observer internal noise variance is also allowed to scale from  $\sigma_e^2$  to  $c^2\sigma_e^2$ ; also note that for an observer with no internal noise, in which we are ultimately interested here, the observer performance  $P_C$ , and Eqn. 18 in general, remain unchanged under such a scale change.)

Finally, Abbey *et al.* define the estimator

$$\hat{\vec{t}} \equiv (\Sigma^+ + \Sigma^-)^{-1} \left( \hat{\vec{z}} - (2\hat{P}_C - 1)\vec{s} \right), \quad (19)$$

where

$$\hat{\vec{z}} \equiv \frac{1}{N_T} \sum_{i=1}^{N_T} \left( \vec{g}_i^{\text{chosen}} - \vec{g}_i^{\text{rejected}} \right) \quad (20)$$

is an unbiased estimator of  $\vec{z}$ , and

$$\hat{P}_C \equiv \frac{N_C}{N_T} \quad (21)$$

is an unbiased estimator of  $P_C$ . ( $N_C$  is the number of correct observer responses in the 2AFC experiment described by Eqn. 20.) Because  $\hat{\vec{z}}$  and  $\hat{P}_C$  are unbiased estimators, as Abbey *et al.* point out,  $\hat{\vec{t}}$  in Eqn. 19 is an unbiased estimator of the scaled observer template  $\alpha\vec{t}$ .

### 2.1.2. Extension to MPMVN case

We define the PDF  $p(\vec{x})$  of a random vector  $\vec{x}$  to be MPMVN with  $N$  peaks if

$$p(\vec{x}) = \sum_{i=1}^N w_i \text{MVN}(\vec{\mu}_i, \Sigma_i), \quad (22)$$

a weighted sum of MVN functions with weights  $w_i$ . Clearly  $\sum_{i=1}^N w_i = 1$  is required for proper normalization. One possible interpretation of this function is as a sum over conditional densities,

$$p(\vec{x}) = \sum_{i=1}^N p(\vec{x}|k_i)p(k_i), \quad (23)$$

by introducing the discrete random variable  $\mathbf{k}$  with  $p(k_i) = w_i$  and  $p(\vec{x}|k_i) = \text{MVN}(\vec{\mu}_i, \Sigma_i)$ , for  $1 \leq i \leq N$ .

We will refer to the parameters  $\vec{\mu}_i$  and  $\Sigma_i$  as the ‘‘component peak’’ means and covariance matrices, respectively. The mean of  $p(\vec{x})$  is easily determined using the linearity of the expectation value operator:

$$\begin{aligned} \vec{\mu} &= \langle \vec{x} \rangle \\ &= \sum_{i=1}^N w_i \langle \vec{x} \rangle_i \\ &= \sum_{i=1}^N w_i \vec{\mu}_i. \end{aligned} \quad (24)$$

Here we have introduced the notation  $\langle \cdot \rangle_i$ , indicating the integral of its argument multiplied by a kernel function which is just the  $i$ th MVN component peak. (Formally, this is equal to the expectation value of a function of a random variable which is itself  $\text{MVN}(\vec{\mu}_i, \Sigma_i)$ .)

The covariance matrix of  $\vec{x}$  can be found using the relation

$$\begin{aligned}
\langle (\vec{x} - \vec{a})(\vec{x} - \vec{a})^\dagger \rangle &= \langle [(\vec{x} - \vec{\mu}) + (\vec{\mu} - \vec{a})][(\vec{x} - \vec{\mu}) + (\vec{\mu} - \vec{a})]^\dagger \rangle \\
&= \langle (\vec{x} - \vec{\mu})(\vec{x} - \vec{\mu})^\dagger \rangle + (\vec{\mu} - \vec{a})(\vec{\mu} - \vec{a})^\dagger \\
&= \Sigma + (\vec{\mu} - \vec{a})(\vec{\mu} - \vec{a})^\dagger,
\end{aligned} \tag{25}$$

where  $\Sigma$  is the covariance matrix of  $\vec{x}$ ,  $\vec{\mu}$  is its mean, and  $\vec{a}$  is any nonrandom vector. Working as in Eqn. 24, we have for an MPMVN random variable

$$\begin{aligned}
\Sigma &= \langle (\vec{x} - \vec{\mu})(\vec{x} - \vec{\mu})^\dagger \rangle \\
&= \sum_{i=1}^N w_i \langle (\vec{x} - \vec{\mu})(\vec{x} - \vec{\mu})^\dagger \rangle_i \\
&= \sum_{i=1}^N w_i [\Sigma_i + (\vec{\mu}_i - \vec{\mu})(\vec{\mu}_i - \vec{\mu})^\dagger].
\end{aligned} \tag{26}$$

Another important property of MPMVN distributions is that if  $\vec{x}$  is MPMVN with  $N_x$  peaks and parameters  $\vec{\mu}_{(x)i}$ ,  $\Sigma_{(x)i}$ , and if  $\vec{y}$  is MPMVN, independent of  $\vec{x}$ , with  $N_y$  peaks and parameters  $\vec{\mu}_{(y)j}$ ,  $\Sigma_{(y)j}$ , then the random variable  $\vec{x} - \vec{y}$  is MPMVN with no more than  $N_x N_y$  peaks:

$$\begin{aligned}
p_{\vec{x}-\vec{y}}(\vec{x} - \vec{y}) &= p_{\vec{x}}(\vec{x}) \star p_{\vec{y}}(\vec{y}) \\
&= \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} w_i w_j \text{MVN}(\vec{\mu}_{(x)i}, \Sigma_{(x)i}) \star \text{MVN}(\vec{\mu}_{(y)j}, \Sigma_{(y)j}) \\
&= \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} w_i w_j \text{MVN}(\vec{\mu}_{(x)i} - \vec{\mu}_{(y)j}, \Sigma_{(x)i} + \Sigma_{(y)j}).
\end{aligned} \tag{27}$$

(Here the symbol  $\star$  indicates the cross correlation operation.)

We now take the same starting point as Abbey *et al.*, namely Eqns. 3, 4, 5, and 10. However, instead of the image density functions given in Eqns. 1 and 2, we draw the images from MPMVN distributions:

$$\vec{g}^+ \sim \sum_{i=1}^{N_+} w_i^+ \text{MVN}(\vec{\mu}_i^+, \Sigma_i^+) \quad \text{and} \tag{28}$$

$$\vec{g}^- \sim \sum_{j=1}^{N_-} w_j^- \text{MVN}(\vec{\mu}_j^-, \Sigma_j^-). \tag{29}$$

Eqn. 13 is unchanged, but instead of Eqn. 12 we have, from Eqn. 27,

$$\Delta \vec{g} \sim \sum_{i=1}^{N_+} \sum_{j=1}^{N_-} w_i^+ w_j^- \text{MVN}(\vec{\mu}_i^+ - \vec{\mu}_j^-, \Sigma_i^+ + \Sigma_j^-). \tag{30}$$

(Note that the signal-present and signal-absent distributions are not required to have the same number of peaks.) We define the signal of interest to be

$$\begin{aligned}
\vec{s} &\equiv \langle \vec{g}^+ \rangle - \langle \vec{g}^- \rangle \\
&= \langle \Delta \vec{g} \rangle \\
&= \sum_{i=1}^{N_+} \sum_{j=1}^{N_-} w_i^+ w_j^- (\vec{\mu}_i^+ - \vec{\mu}_j^-).
\end{aligned} \tag{31}$$

(This is just the generalization of the single-peak case  $\vec{\mu}_1^+ = \vec{s} + \vec{b}$  and  $\vec{\mu}_1^- = \vec{b}$ , as in Eqns. 1 and 2.)

Using these definitions, we can evaluate the expectation value of Eqn. 10:

$$\begin{aligned}
\langle \vec{g}^{\text{chosen}} - \vec{g}^{\text{rejected}} \rangle &= \sum_{i=1}^{N_+} \sum_{j=1}^{N_-} w_i^+ w_j^- \langle \vec{g}^{\text{chosen}} - \vec{g}^{\text{rejected}} \rangle_{ij} \\
&= \sum_{i=1}^{N_+} \sum_{j=1}^{N_-} w_i^+ w_j^- \{ [2(P_C)_{ij} - 1] (\vec{\mu}_i^+ - \vec{\mu}_j^-) + \alpha_{ij} (\Sigma_i^+ + \Sigma_j^-) \vec{t} \} \\
&= \sum_{i=1}^{N_+} \sum_{j=1}^{N_-} w_i^+ w_j^- [2(P_C)_{ij} (\vec{\mu}_i^+ - \vec{\mu}_j^-) + \alpha_{ij} (\Sigma_i^+ + \Sigma_j^-) \vec{t}] - \vec{s},
\end{aligned} \tag{32}$$

where

$$(P_C)_{ij} \equiv \Phi \left( \frac{\vec{t}^\dagger (\vec{\mu}_i^+ - \vec{\mu}_j^-)}{\sqrt{2\sigma_\epsilon^2 + \vec{t}^\dagger (\Sigma_i^+ + \Sigma_j^-) \vec{t}}} \right) \tag{33}$$

and

$$\alpha_{ij} \equiv \frac{2}{\sqrt{2\pi(2\sigma_\epsilon^2 + \vec{t}^\dagger (\Sigma_i^+ + \Sigma_j^-) \vec{t})}} \exp \left( \frac{-(\vec{t}^\dagger (\vec{\mu}_i^+ - \vec{\mu}_j^-))^2}{2(2\sigma_\epsilon^2 + \vec{t}^\dagger (\Sigma_i^+ + \Sigma_j^-) \vec{t})} \right). \tag{34}$$

Eqn. 32 presents us with two difficulties. The first is that we cannot express the right hand side in terms of the observer performance  $P_C$ , as in Eqn. 14. Although it can be shown that

$$P_C = \sum_{i=1}^{N_+} \sum_{j=1}^{N_-} w_i^+ w_j^- (P_C)_{ij}, \tag{35}$$

the expression

$$\sum_{i=1}^{N_+} \sum_{j=1}^{N_-} w_i^+ w_j^- (P_C)_{ij} (\vec{\mu}_i^+ - \vec{\mu}_j^-) \tag{36}$$

in Eqn. 32 cannot be reduced in terms of  $P_C$  for  $\vec{\mu}_i^+$  different from each other and  $\vec{\mu}_j^-$  different from each other. To use Eqn. 32 to construct an estimate of the observer template, we will need to estimate each of the ‘‘component’’ performances  $(P_C)_{ij}$ , by performing independent 2AFC experiments using the same observer but different input images (specifically, images such that  $\Delta \vec{g} \sim \text{MVN}(\vec{\mu}_i^+ - \vec{\mu}_j^-, \Sigma_i^+ + \Sigma_j^-)$ ).

The second difficulty is that the observer template  $\vec{t}$  in Eqn. 32 is not obtainable by a simple matrix inversion. We now have a linear combination of matrices with in principle unknown weights  $\alpha_{ij}$ . If all of these weights were equal, we would be able to invert the sum of known matrices  $(\Sigma_i^+, \Sigma_j^-)$ ; but since the  $\alpha_{ij}$  depend on the unknown template  $\vec{t}$  and the generally different quantities  $(\vec{\mu}_i^+ - \vec{\mu}_j^-)$ , this assumption is not justifiable. The alternative is to restrict the input images to PDFs such that all of the  $\Sigma_i^+$  are equal to a single matrix,  $\Sigma^+$ , and all of the  $\Sigma_j^-$  are equal to a single (in general different) matrix  $\Sigma^-$ . This assumption reduces the generality of our extension of the method of Abbey *et al.*, but appears necessary for the tractability of the problem.

We thus simplify the definitions 28, 29, and 30 to obtain

$$\vec{g}^+ \sim \sum_{i=1}^{N_+} w_i^+ \text{MVN}(\vec{\mu}_i^+, \Sigma^+), \tag{37}$$

$$\vec{g}^- \sim \sum_{j=1}^{N_-} w_j^- \text{MVN}(\vec{\mu}_j^-, \Sigma^-), \quad \text{and} \tag{38}$$

$$\Delta \vec{g} \sim \sum_{i=1}^{N_+} \sum_{j=1}^{N_-} w_i^+ w_j^- \text{MVN}(\vec{\mu}_i^+ - \vec{\mu}_j^-, \Sigma^+ + \Sigma^-). \tag{39}$$

Eqn. 32 now becomes

$$\begin{aligned}
\langle \vec{g}^{\text{chosen}} - \vec{g}^{\text{rejected}} \rangle &= \sum_{i=1}^{N_+} \sum_{j=1}^{N_-} w_i^+ w_j^- [2(P_C)_{ij}(\vec{\mu}_i^+ - \vec{\mu}_j^-) + \alpha_{ij}(\Sigma^+ + \Sigma^-)\vec{t}] - \vec{s} \\
&= \left\{ 2 \sum_{i=1}^{N_+} \sum_{j=1}^{N_-} w_i^+ w_j^- (P_C)_{ij}(\vec{\mu}_i^+ - \vec{\mu}_j^-) \right\} - \vec{s} + (\Sigma^+ + \Sigma^-) \left[ \sum_{i=1}^{N_+} \sum_{j=1}^{N_-} w_i^+ w_j^- \alpha_{ij} \right] \vec{t} \\
&= \vec{s}' - \vec{s} + (\Sigma^+ + \Sigma^-) \alpha_{\text{tot}} \vec{t},
\end{aligned} \tag{40}$$

where

$$\vec{s}' \equiv 2 \sum_{i=1}^{N_+} \sum_{j=1}^{N_-} w_i^+ w_j^- (P_C)_{ij}(\vec{\mu}_i^+ - \vec{\mu}_j^-) \tag{41}$$

and

$$\alpha_{\text{tot}} \equiv \sum_{i=1}^{N_+} \sum_{j=1}^{N_-} w_i^+ w_j^- \alpha_{ij}. \tag{42}$$

Recalling the definition (Eqn. 16) of  $\vec{z}$ , we can rewrite Eqn. 40 as

$$(\Sigma^+ + \Sigma^-)^{-1} [\vec{z} - (\vec{s}' - \vec{s})] = \alpha_{\text{tot}} \vec{t}. \tag{43}$$

Note that in the single peak case,  $\vec{s}'$  reduces to  $2P_C\vec{s}$ ,  $\alpha_{\text{tot}}$  reduces to  $\alpha$ , and we recover Eqn. 18.

We again make use of Eqn. 20 to define an estimator of  $\vec{z}$ , and we define the estimators

$$\hat{\vec{s}}' \equiv 2 \sum_{i=1}^{N_+} \sum_{j=1}^{N_-} w_i^+ w_j^- (\widehat{P_C})_{ij}(\vec{\mu}_i^+ - \vec{\mu}_j^-) \quad \text{and} \tag{44}$$

$$(\widehat{P_C})_{ij} \equiv \frac{(N_C)_{ij}}{w_i^+ w_j^- N_T}, \tag{45}$$

where  $(N_C)_{ij}$  is the number of correctly chosen images in a 2AFC experiment with the same observer but with  $w_i^+ w_j^- N_T$  image pairs drawn from  $\Delta\vec{g} \sim \text{MVN}(\vec{\mu}_i^+ - \vec{\mu}_j^-, \Sigma^+ + \Sigma^-)$ .

We now claim that, given Eqn. 43,

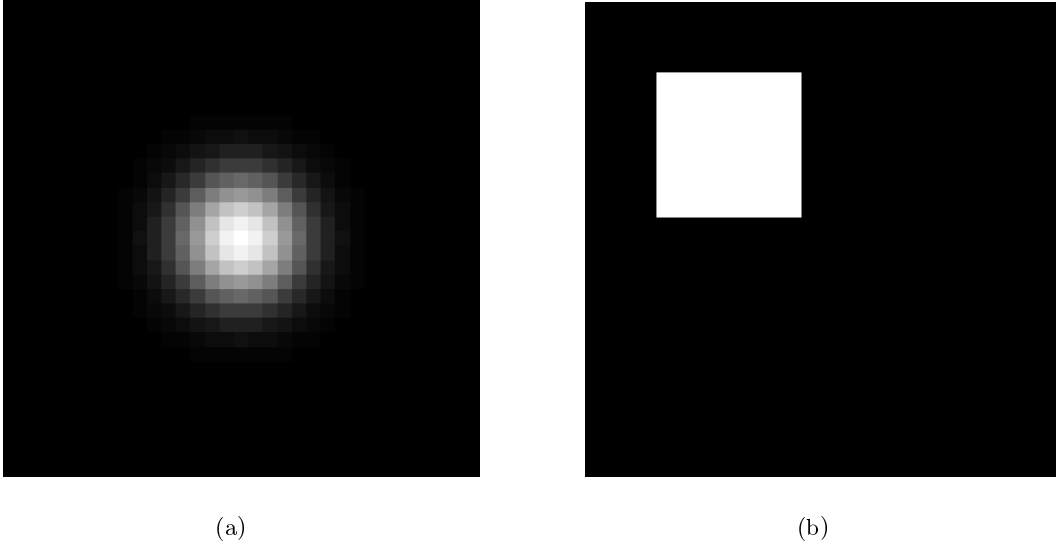
$$\hat{\vec{t}} \equiv (\Sigma^+ + \Sigma^-)^{-1} [\hat{\vec{z}} - (\hat{\vec{s}}' - \vec{s})] \tag{46}$$

is an unbiased estimator of  $\alpha_{\text{tot}}\vec{t}$ , because  $\hat{\vec{z}}$  and  $\hat{\vec{s}}'$  are unbiased estimators of  $\vec{z}$  and  $\vec{s}'$ , respectively.

## 2.2. Simulation studies

For our simulation studies, we used a signal of interest  $\vec{s}$  representing a  $33 \times 33$  pixel Gaussian profile with a center amplitude of 10 graylevels (GL) and a standard deviation of 3 pixels, shown in Fig. 1. The nonrandom background  $\vec{b}$  was 50 GL in each pixel. Two ‘‘mean offset’’ images,  $\vec{n}^+$  and  $\vec{n}^-$ , were generated by pseudorandomly sampling GL values uniformly on the interval  $[0,50]$ . We chose MPMVN parameters such that  $\vec{w}^+ = \vec{w}^- = [0.2, 0.6, 0.2]^\dagger$ , and

$$\begin{aligned}
\vec{\mu}_1^+ &= \vec{s} + \vec{b} - \vec{n}^+ \\
\vec{\mu}_2^+ &= \vec{s} + \vec{b} \\
\vec{\mu}_3^+ &= \vec{s} + \vec{b} + \vec{n}^+ \\
\vec{\mu}_1^- &= \vec{b} - \vec{n}^- \\
\vec{\mu}_2^- &= \vec{b} \\
\vec{\mu}_3^- &= \vec{b} + \vec{n}^-.
\end{aligned} \tag{47}$$



**Figure 1.** (a) Signal of interest  $\vec{s}$ , equal to scaled observer template  $\vec{t}$ . (b) Offset square observer template,  $\vec{t}_{\text{sq}}$ .

Note that  $\sum_{i=1}^3 w_i \vec{\mu}_i^+ = \vec{s} + \vec{b}$  and  $\sum_{j=1}^3 w_j \vec{\mu}_j^- = \vec{b}$  as required.

The covariance matrices  $\Sigma^+$  and  $\Sigma^-$  were chosen equal to one another, with pixel standard deviations of 100 GL. In order to represent a complicated covariance structure, the pixel correlation coefficients were pseudorandomly generated uniformly and then corrected to make the matrix a valid covariance matrix (symmetric with positive definite eigenvalues).

Two observer templates were studied, shown in Fig. 1. The first was simply a normalized version of the signal,  $\vec{t} \equiv \vec{s}/\sqrt{\vec{s}^\dagger \vec{s}}$ ; we will refer to this as the signal-equivalent template. The second was an offset square mask  $\vec{t}_{\text{sq}}$ , with values of 1 in pixel locations (6,6) through (15,15) and 0 elsewhere. The observer was deterministic, and all relations developed in Section 2.1.2 were used with the internal noise variance  $\sigma_\epsilon^2$  set to 0.

In each simulation study, a value of  $N_T$  was selected, and  $N_T$  pairs of images were generated pseudorandomly from the distributions in Eqns. 37 and 38 using the parameters stated above. A 2AFC experiment was then performed to produce the estimator  $\hat{\vec{z}}$ . An independent set of  $N_T$  image pairs was again generated to produce  $\hat{\vec{z}}'$ . The estimators were combined using Eqn. 46 to produce the template estimate  $\hat{\vec{t}}$ . We then calculated the mean squared error (MSE) of this estimate, where

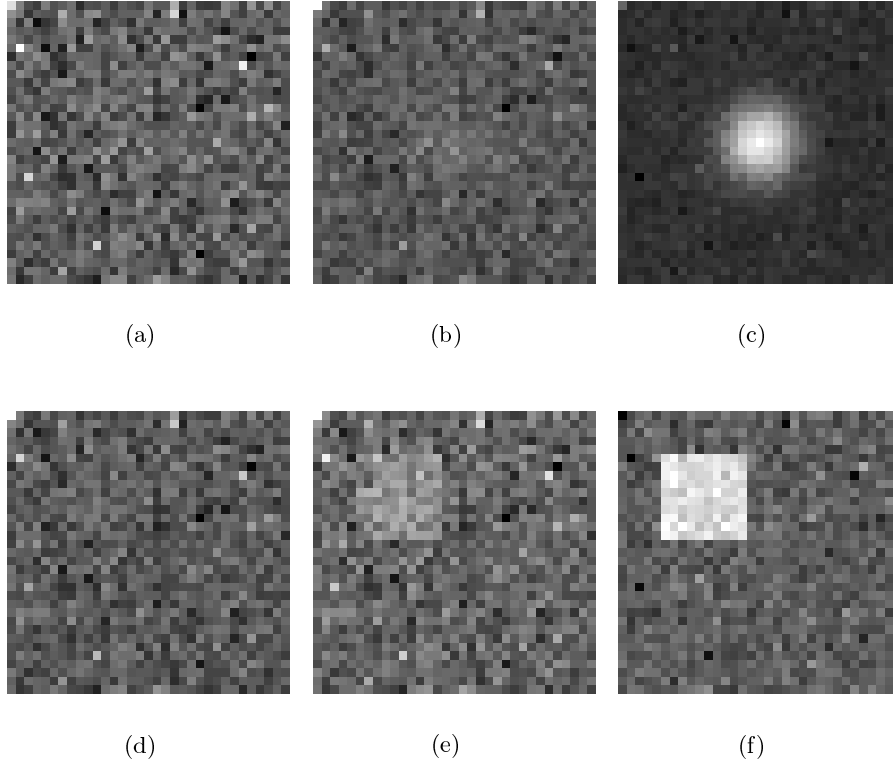
$$\mathbf{MSE} \equiv (\hat{\vec{t}} - \alpha_{\text{tot}} \vec{t})^\dagger (\hat{\vec{t}} - \alpha_{\text{tot}} \vec{t}). \quad (48)$$

Values of  $N_T$  used were 10,000, 20,000, 50,000, 100,000, 200,000, 500,000, 1,000,000, 5,000,000, and 20,000,000.

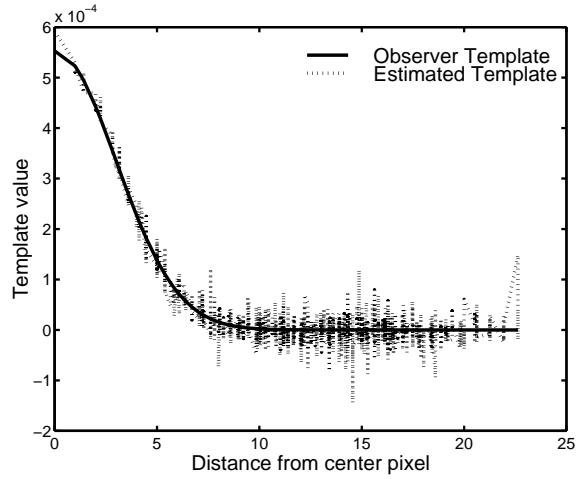
We also measured the reliability of the MSE measure by generating 100 template estimates, each with  $N_T = 100,000$ . The sample mean  $\widehat{\mathbf{MSE}}$  of the MSE values was calculated, as an estimate of the expectation value of the MSE. This expectation value can be calculated explicitly:

$$\begin{aligned} \langle \mathbf{MSE} \rangle &= \left\langle (\hat{\vec{t}} - \alpha_{\text{tot}} \vec{t})^\dagger (\hat{\vec{t}} - \alpha_{\text{tot}} \vec{t}) \right\rangle \\ &= \left\langle \text{Tr} \left[ (\hat{\vec{t}} - \alpha_{\text{tot}} \vec{t}) (\hat{\vec{t}} - \alpha_{\text{tot}} \vec{t})^\dagger \right] \right\rangle \\ &= \text{Tr} \left( \left\langle (\hat{\vec{t}} - \alpha_{\text{tot}} \vec{t}) (\hat{\vec{t}} - \alpha_{\text{tot}} \vec{t})^\dagger \right\rangle \right) \\ &= \text{Tr} (\Sigma_t), \end{aligned} \quad (49)$$

the trace of the covariance matrix of the template estimate  $\hat{\vec{t}}$  (for a given number of 2AFC image pairs  $N_T$ ). We compared the calculated trace with the measured sample mean  $\widehat{\mathbf{MSE}}$ .



**Figure 2.** Estimates of  $\vec{t}$  for  $N_T =$  (a) 50,000, (b) 1,000,000, and (c) 5,000,000; and estimates of  $\vec{t}_{\text{sq}}$  for  $N_T =$  (d) 50,000, (e) 1,000,000, and (f) 5,000,000.

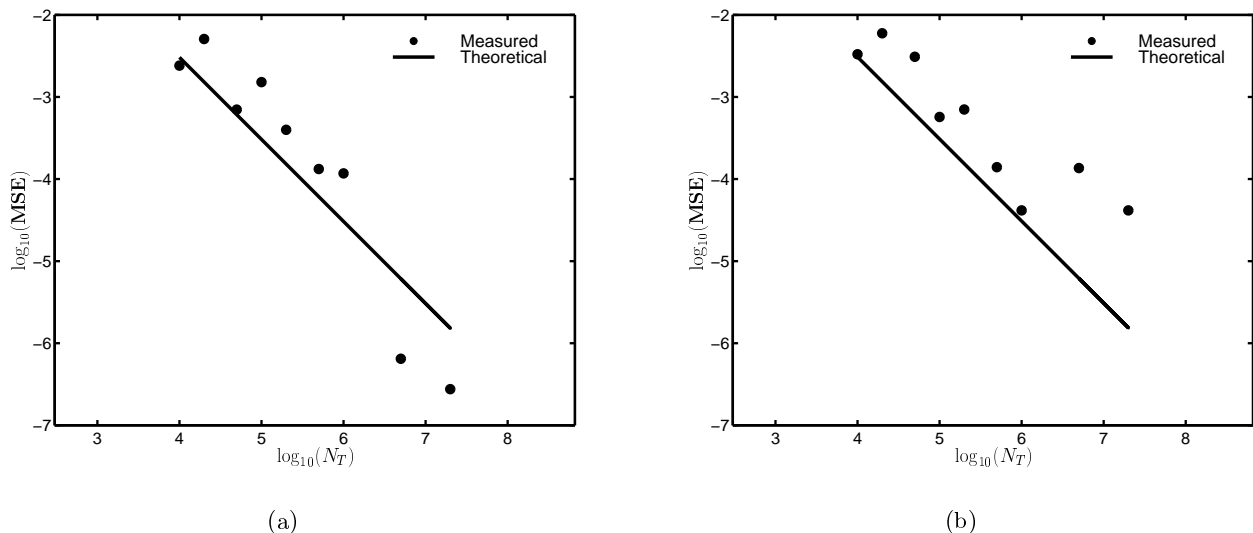


**Figure 3.** Signal-equivalent template estimate,  $N_T = 5,000,000$ , and the actual observer template.

### 3. RESULTS

Template estimates for the signal-equivalent and square observer templates are shown in Fig. 2. Estimates for three values of  $N_T$  are given for each observer template.

The signal-equivalent template estimate for  $N_T = 5,000,000$  is shown with the observer template in Fig. 3 as a function of distance from the image center in pixels. Graphs of  $\mathbf{MSE}$  and  $\text{Tr}(\Sigma_t)$  as functions of  $N_T$  for both observer templates are shown in Fig. 4.



**Figure 4.** Graphs of **MSE** and the covariance matrix of the template estimate *vs.*  $N_T$  for: (a) signal-equivalent template  $\vec{t}$ , (b) square template  $\vec{t}_{sq}$ .

The trace of the covariance matrix of the template estimate, for  $\vec{t}_{sq}$  and  $N_T = 100,000$ , was  $3.08 \times 10^{-4}$ , which compares reasonably well with the sample mean  $\widehat{\mathbf{MSE}} = 2.47 \times 10^{-4}$  for 100 sample template estimates. For the signal-equivalent template, with the same numbers of image pairs and template estimates, the trace of the covariance matrix was  $3.05 \times 10^{-4}$ , while the sample mean was  $\widehat{\mathbf{MSE}} = 2.93 \times 10^{-4}$ .

#### 4. DISCUSSION

The template estimates produced by our extension of the method of Abbey *et al.* are evidently unbiased (Figs. 2 and 3). Fig. 4 shows that the covariance matrix of the template estimate is inversely proportional to the number of image pair samples, as might be expected. (The variance of a sample mean is inversely proportional to the number of samples, and the template estimate in Eqn. 46 can be interpreted as a linear function of sample means  $\hat{\vec{z}}$  and  $(\widehat{\mathbf{P}\mathbf{C}})_{ij}$ .) However, it is clear from this graph that the values of **MSE** do not agree very well with the traces of the covariance matrices for smaller values of  $N_T$ . This is not particularly surprising, since these values of **MSE** are effectively “sample means” over a single sample, *i. e.*, over a single template estimate. Because the template estimate itself can be so readily interpreted in terms of sample means of image pair properties, an estimate of  $\langle \mathbf{MSE} \rangle$  over a single template estimate obtained with a very large number of image pairs  $N_T$  is equivalent to an estimate of  $\langle \mathbf{MSE} \rangle$  over many template estimates, each obtained with proportionally fewer image pairs. Thus, the agreement between **MSE** and the trace of the covariance matrix of the template estimate is much better for higher values of  $N_T$ .

Comparison of the images in Fig. 2 with those presented in Ref. 5 indicates that many more 2AFC image pairs are required in the MSMVN case than in the single peak MVN case to obtain template estimates of comparable quality. By calculating the template estimate covariance matrices for both cases, one can predict the increase in  $N_T$  required. It turns out however that the actual increase in the required number of image pairs is much larger than that predicted by the template estimate covariance matrices alone. The reason for this is still under investigation.

#### 5. CONCLUSIONS

We have shown that the method of Abbey *et al.* for constructing an unbiased estimate of a linear observer template using the results of a 2AFC experiment with MVN-noise images<sup>5</sup> can be extended to 2AFC images in which the pixel noise is drawn from an MPMVN distribution, *i. e.*, a weighted sum of MVN distributions.

For an MPMVN distribution with given component peak covariance matrices, the extension requires many more 2AFC image pairs for a good template estimate than would the original method for images with a noise covariance

matrix equal to that of a single peak of the MPMVN distribution. One reason for this is the corresponding increase in the template estimate covariance matrix, which depends upon not only the “widths” of the component peaks (the covariance matrices), but also the distances between them.

Another complication of the extension relative to the original method is the requirement of a second set of  $N_T$  images, with noise properties determined only by individual peaks of the MPMVN distribution, in order to obtain an estimate of  $\hat{\mathbf{z}}'$ , a component of the template estimate. For simulation studies such as those presented here, this is not a problem. In experiments using actual images whose pixel noise distributions are being approximated by an MPMVN distribution, the accuracy of the template estimate will clearly depend on the quality of this approximation and on the feasibility of producing simulated images containing only component peak noise.

Our primary goal here has been to demonstrate that the method of Abbey *et al.* can be extended in nontrivial ways. We expect that the original method, as well as extensions such as that presented here, will yield many more interesting results on further investigation.

## ACKNOWLEDGMENTS

The first author thanks Dr. Robert Wagner for bringing Ref. 5 to his attention, and for helpful discussions on its contents.

## REFERENCES

1. K. Doi, “Computer-aided diagnosis and its potential impact on diagnostic radiology,” in *Computer-Aided Diagnosis in Medical Imaging*, K. Doi, H. MacMahon, M. L. Giger, and K. R. Hoffmann, eds., Proceedings of the First International Workshop on Computer-Aided Diagnosis, pp. 11–20, Elsevier, (Amsterdam), 1999.
2. J. Yerushalmy, J. T. Harkness, J. H. Hope, and B. R. Kennedy, “The role of dual reading in mass radiography,” *American Review of Tuberculosis* **61**, pp. 443–464, 1950.
3. B. J. Hillman, S. J. Hessel, R. G. Swensson, and P. G. Herman, “Improving diagnostic accuracy: A comparison of interactive and delphi consultations,” *Investigative Radiology* **2**, pp. 112–115, 1977.
4. S. J. Hessel, P. G. Herman, and R. G. Swensson, “Improving performance by multiple interpretations of chest radiographs: Effectiveness and cost,” *Diagnostic Radiology* **127**, pp. 589–594, 1978.
5. C. K. Abbey, M. P. Eckstein, and F. O. Bochud, “Estimation of human-observer templates in two-alternative forced-choice experiments,” in *Medical Imaging 1999: Image Perception and Performance*, E. A. Krupinski, ed., *Proceedings of SPIE* **Vol. 3663**, pp. 284–295, 1999.